2015

Detection and Characterization of Pathogenic Mycobacteria Using Binary Deoxyribozymes

Bradley Rosenkrantz
University of Central Florida

DETECTION AND CHARACTERIZATION OF PATHOGENIC MYCOBACTERIA USING BINARY DEOXYRIBOZYMES

by

BRADLEY A. ROSENKRANTZ

A thesis submitted in partial fulfillment of the requirements
for the Honors in the Major Program in Biomedical Sciences
in the Burnett School of Biomedical Sciences
and in The Burnett Honors College
at the University of Central Florida
Orlando, Florida

Spring Term 2015

Thesis Chair: Dr. Kyle Rohde

Abstract

The genus Mycobacterium contains many pathogenic bacteria that are known to cause serious diseases in humans. One of the most well-known of these bacteria is *Mycobacterium tuberculosis*, or *Mtb*, which is the causative agent of tuberculosis. It infects nearly one-third of the world's population and kills 1.4 million people annually. Another important mycobacterial pathogen is *Mycobacterium abscessus*, or *Mabs*, which causes respiratory infections in cystic fibrosis patients. One of the biggest difficulties in combating these pathogens is the lack of effective diagnostics, as current strategies hold many pitfalls and can be unreliable. One common method used is sputum smear microscopy which involves acid fast staining of the bacteria present in a patient's sputum. This method of detection fails to detect more than 50% of infections and is unable to differentiate between species of mycobacterium. This project introduces a novel method of mycobacterial diagnostics using binary deoxyribozymes (DNAzymes). Binary DNAzymes recognize bacteria-specific nucleic acid sequences and bind to them, forming a catalytic core which cleaves a substrate molecule. This cleavage separates a quencher molecule from a fluorophore, which results in a fluorescent output. This flexible assay platform has great potential for the detection of *Mtb* or *Mabs*. Our data shows the specificity of the DNAzymes allowing for a differential diagnosis of various species of Mycobacteria. It also shows the limit of detection of this technology and its additional utility in molecular typing of *Mtb* clinical isolates as well as drug resistance characterization. This multipurpose tool can contribute to disease management in multiple ways.

Acknowledgments

I would like to thank both Dr. Rohde and Dr. Kolpashchikov for the guidance in making this thesis possible. I would also like to thank both the members of the Rohde and Kolpashchikov labs for their continued support. Finally I would like to thank Denise Crisafi for convincing me that writing a thesis was actually a good idea.

Dedication

I dedicate this body of work to the 308 family
Reach for the stars so if you fall, you land on a cloud

Table of Contents

List of Figures

List of Tables

Introduction

Pathogenic Mycobacteria

Mycobacterium tuberculosis

The genus Mycobacterium contains many pathogenic bacteria that are known to cause serious diseases in humans. One of the most well-known of these bacteria is *Mycobacterium tuberculosis* (*Mtb)*, which is the causative agent of tuberculosis (TB). *Mtb* infects nearly one-third of the world's population and kills 1.4 million people annually [1]. Although *Mtb* was first identified as the causative agent of TB in 1893 by Robert Koch, *Mtb*-human interactions date back to ancient Egypt and even earlier [2, 3].

Mtb is an acid-fast staining gram-positive bacillus, due to its cell wall's ability to take up a positively charged dye. This characteristic is indicative of the molecular structure of the cell wall, which is composed primarily of mycolic acid and mycolic acid derivatives [4]. These lipid chains vary in size, ranging from 60 to 90 carbons long [5]. The specific lipids may vary between different species in the *Mycobacterium* genus [4]. These long chain fatty acids can elicit an immune response in a mouse animal model, indicating their recognition as a PAMP, or Pathogen Associated Molecular Pattern, by the host immune system [6]. The mycolic acid along with the LPS layers form an envelope which protects the bacteria from host defenses such as degradative enzymes by forming a permeability barrier and free radicals via its cyclopropionated mycolic acid residues [7].

Mtb is transmitted via aerosolized droplets produced when an infected individual coughs and expels the bacteria into the environment. Transmission is accomplished when the bacteria is taken up into a host's airways and into the lungs where it is internalized by alveolar macrophages

into a phagosome [8]. While in the phagosome the bacteria secrete effector proteins which manipulate the phagosome morphology and recruitment of host proteins involved in vesicle trafficking. This prevents fusion of the lysosome, allowing the bacteria to survive within the phagosome and avoid both acid exposure and hydrolytic enzymes [9]. In the case of active TB disease, which only occurs in 5 – 10% of *Mtb* infections, the bacteria are able to multiply unchecked within the macrophage and eventually cause cell necrosis and tissue damage [10]. It is from this point that the bacteria may spread into the capillaries of the lungs and move systemically throughout the host. Transmission is accomplished due to the coughing response to the necrosis within the lungs [8].

The hallmark of a pulmonary TB infection is the presence of a granuloma, a tumor-like walled off nodule containing infected macrophages, neutrophils, and dendritic cells [8]. Within the granuloma, the *Mtb* are located within the macrophages directly within the center of the granuloma [11]. These infected cells are surrounded by abnormal immune cells such as giant multinucleated cells and fat containing "foamy macrophages [12]." The formation of a granuloma may be an innate immune response to prevent transmission, however current theories exist that the granuloma formation is evolutionarily beneficial to the bacteria by offering both protection and nutrients [13]. *Mtb* within the granuloma are dormant, replicating infrequently and become highly drug resistant via intrinsic biochemical mechanisms such as efflux pumps [8, 14]. About 90% of those infected with *Mtb* do not display symptoms immediately and are referred to as having a latent TB infection (LTBI). In a LTBI, *Mtb* growth is restrained by activated macrophages, limiting their ability to spread. An acute production of T-cells is seen at the peripheral lymph nodes, however production of naïve T cells decline during the chronic

infection due to inefficient T cell priming [15]. Th1 T-cells produced aid in preventing reactivation of the granuloma via the production of cytokines, however research suggest some cytokines may activate *Mtb* dormancy genes [16]. LTBI can remain asymptomatic for years or in 5-10% of the cases can reactivate to produce a symptomatic disease. Some research suggests now that the recruited neutrophils may play a role in the progression of latent to active Tb infection, causing necrotic cell death of the macrophages within the granuloma containing the *Mtb* [8]. This results in caseous necrosis of the granuloma, which facilitates escape and transmission of the bacteria and allows the cycle to repeat [8, 14] .

Mycobacterium abscessus

Another pathogenic member of the Mycobacterium genus is *Mycobacterium abscessus* (*Mabs).* *Mabs* is a rapidly growing acid-fast gram-positive rod. *Mabs* acid-fast properties, similar to *Mtb*, are due to the high concentration of mycolic acid present in its cell wall [4]. *Mabs* was first described in 1953 after it was identified from an abscess in a patients knee, and was originally classified as a subspecies of *M. chelonae* [17]. In 1992 genomic analysis and hybridization studies allowed the determination that *Mabs* is sufficiently distinct enough from *Mycobacterium chelonae* to warrant reclassification as a distinct species [18].

Mabs infections are commonly seen in patients who are immunocompromised. *Mabs* has therefore been noted as a common cause of hospital borne infections in a patient post-surgery [19-21]. *Mabs* infections can be pulmonary, looking remarkably similar to a *Mtb* infection [22]. Patients with cystic fibrosis are at a significantly higher risk of developing a pulmonary infection from *Mabs* than the general public [23]. Studies estimate that 5-12% of all CF patients are infected with *Mabs*, with that number on the rise [22, 24, 25]. *Mabs* infections can also be

extrapulmonary and can occur anywhere on the host, including the eye, with a majority reported as subcutaneous [26]. An extrapulmonary *Mabs* infection can commonly produce an abscess, or walled off fluid filled sac, giving *Mycobacterium abscessus* its namesake [27]. Although originally considered rare *Mabs* infections are now 10 times more common in the USA than *Mtb* infections [19, 28].

Drug Resistance

Treatment of both *Mtb* and *Mabs* involve the usage of different multi-drug cocktails which are taken for 6-9 months [29, 30]. *Mabs* is extremely difficult to treat due to intrinsic drug resistances to a variety of treatment options, including the drugs used to treat *Mtb.* The effects of its lipid envelope, degradative enzymes and efflux pumps, result in an even longer antibiotic treatment regimen than that for *Mtb* [31, 32]. *Mtb* has been shown to express a variety of mechanisms to survive drug treatment. Dormant bacteria may be signaled from a variety of stress conditions such as the decrease in pH or hypoxia inside of the phagosome to begin transcription of essential genes [33]. This results in the upregulation of multiple efflux pumps and drug inactivating enzymes which confer drug resistance to several antibiotic compounds such as tetracycline [34-37]. The complex regulons which control these efflux pumps are currently new targets of antimicrobial research and development.

Treatment of *Mtb* has become increasingly difficult due to the emergence of drug resistance mutations in the *Mtb* genome. It has been shown that SNP's, or single nucleotide polymorphisms, can alter a drug target and prevent antimicrobials from inhibiting their target's function. *Mtb* can acquire multiple drug resistance mutations, resulting in Multi-Drug Resistant Tuberculosis (MDR-TB). A *Mtb* infection will be diagnosed as MDR-TB if it is shown that the

bacteria has resistance to both isoniazid and rifampin, two first line treatment options [38]. Rifampin (Rif) targets the β subunit of the RNA polymerase, encoded by the *rpoB* gene [39]. It has been shown that the *rpoB* gene contains an 81 base pair hotspot region which has a significantly higher rate of mutations in Rif-resistant strains of *Mtb*, indicating a genetic target for drug-mediated artificial selection [40]. Mutations in both the 526 and 531 codons have commonly been shown to confer drug resistance, altering the Rif binding site on the β subunit [41].

Isoniazid (Inh) is another first line treatment option for *Mtb* [38]. Isoniazid binds the product of the *inhA* gene, a long-chain enoyl-acyl carrier protein reductase involved in elongation of the mycolic acid fatty chains [42, 43]. Mutations in the *inhA* gene confer resistance to Inh by preventing binding in the active site, however the occurrence of a mutation in the ORF (open reading frame) is not commonly the case [44-46]. Mutations within the *inhA* promoter also result in resistance due to increased production of the protein product, resulting in low-level resistance to clinically used Inh concentrations [47]. Missense mutations in the *katG* gene, a gene which encodes a catalase-peroxidase enzyme, also confer Inh resistance [45, 48-50]. The KatG protein activates Inh by coupling it to NAD(+)/ NADH, forming an active product. Mutations in KatG prevent this activity, however both catalase and peroxidase activity is maintained indicating a more complex relationship between KatG and Inh than originally hypothesized [49, 51].

Extensively drug resistant TB (XDR-TB) is a subtype of MDR-TB where the bacteria are additionally resistant to a fluoroquinolone (FQ) as well as an injectable second line drug. Fluoroquinolones bind to the gyrase and topoisomerase enzymes used to move the replication

fork forward during DNA replication and relieve tension by unwinding the DNA [52]. The binding of FQs prevents DNA replication and damages the genome, eliciting an SOS cellular response and the generation of oxygen radicals resulting in cell death [53]. Resistance to FQs is due to mutations in the *gyrA* gene or the *gyrB* gene, both of which encode gyrase enzymes in *Mtb* [54]. These mutations change the amino acid sequence in the FQ binding site, dramatically increasing its minimum inhibitory concentration [55].

Diagnostics and Drug Resistance Characterization

Before the development of molecular techniques, a pathogenic mycobacterial infection was diagnosed based on sputum smear microscopy [56]. This technique is quick and inexpensive, which is why it is still currently utilized as a Point of Care (POC) diagnostic. An infected patient is required to provide a sputum sample, which is then processed to isolate the bacteria. The sample is then stained with carbol-fuchsin and counterstained with methylene blue. Due to the mycolic acid component of the cell wall, the carbol-fuchsin will be retained by the mycobacterium, resulting in a red appearance. This method of staining is also referred to as the Ziehl-Neelsen stain, and can detect a minimum of 5,000 to 10,000 bacilli per milliliter [57, 58]. Acid fast staining has been the gold standard for detecting *Mtb* for over 100 years due to its cost effectiveness. Unfortunately, this technique is highly ineffective at detecting an infection, failing to do so almost 50% of the time [59]. Sensitive detection is even more problematic in children and patients that are co-infected with HIV [60, 61]. This is likely due to the inability to obtain a large enough sputum sample to visualize the bacteria. Higher occurrences of extra pulmonary TB in immunocompromised patients also make diagnostics using sputum more

difficult [60]. The Ziehl-Neelsen stain is also unable to differentiate between different species in the *Mycobacterium* genus.

Another non-molecular method of diagnosing pathogenic mycobacteria is by using an apparatus known as a Mycobacteria growth indicator tube (or MGIT). The MGIT tube contains Middlebrook 7H9 Broth, a liquid growth medium containing the essential composition for Mycobacterial growth [62]. After a patient's sputum is processed, the sample is placed in the MGIT tube. If mycobacteria are present then they will begin to grow in the tube and respire using the dissolved oxygen. The depletion of oxygen in the tube results in the activation of a fluorescent compound and therefore detection. The increasing cases of drug resistant *Mtb* has resulted in resistance characterization becoming a necessary and integral part of both treatment and diagnostics [63]. Adding various drugs such as Rif and Inh to the MGIT allow this test to be utilized for drug susceptibility testing [38].

Another common technique used to diagnose *Mtb* is the Mantoux test, also known as the PPD (Purified Protein Derivative) skin test or the Tuberculin skin test. This test relies on analyzing the body's response to a small injection containing tuberculin, a protein extract from *Mtb* [64]. The actual make-up of PPD may vary and is known to contain many different antigens. After the tuberculin has been injected, the size of the raised skin is measured in millimeters. A positive result indicates the host has already mounted an immune response to *Mtb* or an *Mtb* like antigen. The criteria for diagnostics may vary between patients, with specific size intervals and additional risk factors of the patient contributing to the analysis [64]. A chest X-ray usually follows to determine if an active disease state is present.

The PPD skin test, which does not indicate an active infection but only the presence of a secondary immune response, is known to cross react with patients who have previously been vaccinated with the BCG vaccine. The BCG, or Bacille de Calmette et Guérin, vaccine is made of an attenuated strain of *Mycobacterium bovis*, a mycobacterium which causes TB-like symptoms in cattle. *Mycobacterium bovis* contains enough similarities with *Mtb* to elicit an immune response, however the efficacy of the vaccine is anywhere from 0 – 80%, which is why the vaccine is not widely utilized in the United States (however it is used widely around other parts of the world) [65]. Due to the antigenic similarities between Tuberculin and the BCG vaccine, those previously vaccinated will elicit a more powerful immune response, resulting in a false positive. False positives may also result from an infection by other members of the Mycobacterium genus, such as those found in the environment [66, 67].

Modern molecular biology has opened the door for new methods of diagnostics of pathogenic mycobacteria. Nucleic Acid Amplification Tests (or NAATs), rely on the usage of polymerase chain reaction (PCR) to amplify specific sequences of DNA derived from the bacteria [68]. Common targets for identification of the bacteria are the highly conserved and species specific 16S and 23S rRNA of the ribosome [69]. The amplicon may then be sequenced for diagnosis or hybridized to a molecular probe [70]. A hybridization probe is a synthetic oligonucleotide containing a sequence that is complimentary to the target sequence. Binding of the molecular beacon may separate a fluorophore from a quencher molecule, resulting in a fluorescent signal. By targeting the genes responsible for drug resistances, molecular probes are also able to characterize the drug resistances of the bacteria [71]. This can also be determined via DNA sequencing. New advancement in molecular diagnostics has allowed the process to be

miniaturized by using oligonucleotides covalently linked to a membrane or a microchip [72]. This technique is highly effective and only requires a small amount of DNA.

A commercially available diagnostics technique that implements NAATs is the Cepheid Gene Xpert which can both detect *Mtb* and Rif drug resistance. The WHO has financed promotion of the Gene Xpert, subsidizing the costs of the instrument and assay reagents in resource poor countries with high endemic TB rates. This technology takes two hours to both analyze the presence of *Mtb* and characterize Rif drug resistance [73, 74]. Unprocessed sputum is sonicated to lyse the bacteria and the nucleic acid is purified and concentrated. The nucleic acid is then PCR amplified and hybridized to probes which provide species identification along with probes which target the *rpoB* gene. Gene Xpert's sensitivity was determined via a LOD study to be 131 CFU per ml and can accurately differentiate against non-tuberculosis mycobacteria by targeting nonhomologous regions of the *rpoB* gene [73]. The *rpoB* gene is targeted by 5 separate molecular beacons, each requiring their own unique color to both diagnose and characterize *Mtb* [75]. Unfortunately it seems this technology has only proved cost effective in reference centers, requiring expensive equipment and reagents. The requirement of electricity and trained personnel makes this technique difficult to implement into a Point of Care diagnostic setting [76].

Mtb Genetic Diversity and Classification

Molecular tools for genetic characterization of *Mtb* revealed lineage specific differences, allowing populations to be sub-classified into both lineages and strains [77]. The *Mtb* lineages can be classified as either modern or ancient, with lineages 2-4 considered modern. This classification is based on the presence or absence of TbD1, which is absent in modern strains

[78]. There are 7 *Mtb* lineages that can infect humans, each of which are associated with different geographical regions (see table 1). The seven lineages can be further divided into different strains, each of which has a unique genetic makeup. The two strains of *Mtb* that are most common in research lab settings are H37Rv and CDC1551.

Table 1: *Mtb* Lineages

Lineage	Geographic Classification
1	Indo-Oceanic
2	East-Asian
3	East- African – Indian
4	Euro-American
5	West- Africa
6	West- Africa
7	Ethiopia

Because the different strains are over 90% genetically identical, it was originally believed that there were no significant differences between them. Recent evidence however suggests that the different *Mtb* lineages differ in virulence [77]. Unlike many other pathogenic bacteria, *Mtb* does not possess any toxin mechanisms; therefore virulence is often directly equated to transmission. It has also been seen that a lower immunological response is correlated to higher bacterial populations in the host, resulting in higher transmission[77]. Modern strains are commonly associated with a delayed inflammatory response and are therefore seen to exhibit a higher level of transmission along with differences in global gene expression and macrophage survival [79, 80]. These modern lineages are also associated with a higher likelihood of progression to an active disease state and are seen more globally than the ancient strains, suggesting that the delayed inflammatory response may be associated for the globalization of the lineages. Between the modern lineages, lineage 2 is seen to have the highest virulence along with the highest occurrences of *de novo* drug resistances [81]. Ford *et al*. have shown that this

may be a result of the higher mutation rate seen within lineage 2 [82]. These phenotypic differences illustrate the effect in both the disease progression and treatment that the *Mtb*'s genetic variability has on an infection.

Mtb lineage and strain identification is useful to track *Mtb*'s virulence, evolution, and for epidemiological studies [83, 84]. The spread of *Mtb* can be mapped out according to lineage, allowing determination of population based infection patterns. Analysis of the *Mtb* genome in seals suggests a similarity between the murine strain and the strains seen in the Americas, suggesting that *Mtb* may have jumped from seals to humans [85]. On a micro-population basis, analysis of the different lineages and strains can elucidate infection patterns allowing the determination of transmission of *Mtb* between individual patients. It is also useful in the determination of nosocomial infections along with the origin of MDR/XDR-TB, allowing researchers to analyze whether a patient infected with *Mtb* acquired the infection within the hospital or before entering it [86]. Finally, strain typing can be used in conjunction with drug resistance determination in order to determine the prevalence of *de novo* mutations conferring drug resistance. This can be used to help determine if drug resistant *Mtb* strains in health-care settings are due to internal transmission of the MDR-TB or due to evolution of resistance during treatment selection.

Strain Typing Methods

A variety of molecular techniques have been used to classify and type *Mtb*, analyzing genetic markers such as SNPs, insertions, repeats and deletions. The first molecular technique used to analyze *Mtb* was restriction fragment length polymorphism (RFLP) typing. This technique relies on a difference in copy numbers and location of a genomic insertion sequence

(IS6110) in the *Mtb* genome and was originally hailed as the "gold standard" of strain typing [87]. RFLP involves the digestion of the chromosomal DNA with an endonuclease, resulting in the production of many fragments of variable lengths. The digested DNA is then separated on an agarose gel via electrophoresis, where DNA fragments of larger sizes migrate slower than those of smaller sizes. The gel is then blotted onto a nylon membrane which absorbs the DNA and retains their relative positioning on the gel. The IS6110 insertion sequence is then detected using fluorescent or radioactive probes and imaged. This allows for determination of repeat sequences relative to the migration pattern of the DNA band and therefore characterization of the strains [88, 89]. Unfortunately this technique requires large amount of DNA and has a propensity of contamination, illustrating the need of newer technology.

Only a few years later, the need for more effective molecular techniques led the way to PCR-based methods of analysis being developed. By utilizing primers specific to sequences found only in *Mtb*, the specificity of the PCR based characterization methods were determined to be 100% [77]. Two methods of PCR based analysis that have become common in strain identification include Spoligotyping and MIRU typing. Spoligotyping, or spacer oligotyping, involves PCR amplification of 43 unique regions located between direct repeats in the Clustered Regularly Interspaced Short Palindromic Repeat region (CRISPRs) [90]. This results in the amplification of a large number of fragments which may vary in size. The amplified fragments are then hybridized to a membrane which has the corresponding oligonucleotides covalently bonded. The hybridization pattern differs from each lineage and strain of *Mtb*, allowing for effective strain identification [88]. This technique can be used in conjunction with RFLP typing or independently [91].

MIRU typing differs from spoligotyping in the target of the PCR amplification. MIRU typing is accomplished via the PCR amplification of MIRUs, or *M*ycobacterial *I*nterspersed *R*epetitive *U*nits [92]. These MIRU's are found within the *Mtb* genome and are reminiscent of human Variable Number Tandem Repeats (VNTRs) [93]. The MIRUs of interest are amplified via PCR and separated using gel electrophoresis. The gel is then stained and the size of the band is determined via comparison to a molecular marker of known sizes [89]. Results from both MIRU typing and Spoligotyping can be compared to a known culture or compared to an online database, allowing for strain and lineage determination [94, 95]. Unfortunately, identical patterns between strains can result from convergent evolution which would make strain identification and evolutionary determination nearly impossible [96].

Another method of characterization of the *Mtb* strains and lineage is the determination of Large Sequence Polymorphisms (LSPs), also referred to as Regions of Difference (RDs). This method of characterization relies on amplification of a region of the *Mtb* genome containing a large deletion, followed by gel electrophoresis and detection using probe hybridization [97]. Because large genomic deletions are irreversible, LSP analysis is very useful for determination of a lineages evolutionary history [98].

Spoligotyping, MIRU typing and LSP analysis have revealed a high level of variability within lineages and strains of *Mtb*, along with similarities between strains and lineages which may have resulted from independent evolutionary events. This has resulted in the use of whole genome sequencing in determining the phylogenetic lineage of an *Mtb* strain [99, 100]. Whole genome sequencing was first accomplished via the shotgun sequencing method, where the entire genome is randomly cleaved and each fragment in inserted into a plasmid and transformed into

competent cells. The individual fragments were then sequenced and assembled using a variety of computer programs [101]. Whole genome analysis not only showed LSP but also elucidated a variety of important SNPs, or single nucleotide polymorphisms. These SNPs are differences in a single nucleotide between lineages or strains of *Mtb* [102, 103]. Recently an algorithm has been developed which allows lineage determination based on the sequencing of only a few genes, relying totally on the determination of a few SNPs in order to classify a strain of *Mtb* [104].

Binary Deoxyribozymes

One form of molecular diagnostics and characterization that has not been applied to *Mtb* or *Mabs* before is a binary deoxyribozyme, or DNAzyme. DNAzymes are catalytic strands of DNA based on naturally occurring ribozymes which, in the presence of a metal cofactor, have been shown to facilitate the cleavage of an RNA or RNA/DNA phosphodiester bond [105]. Variations exist however which have peroxidase like activity or silence genes via mRNA cleavage [106, 107]. One of the most studied DNAzymes is the 10-23 DNAzyme which cleaves an RNA molecule between a purine and pyrimidine [108]. The region of the DNAzyme which is enzymatically active is referred to as the catalytic core (see Figure 1a). The deoxyribozyme was first developed by Breaker and Joyce after selective amplification of random nucleotides [109, 110]. The use of a peptide bound oligonucleotide allowed for the *in vitro* selection of catalytically active DNA sequences [111].

UCF professor and collaborator Dr. Dmitry Kolpashchikov, as well as others, have shown that a variety of modifications can transform the DNAzyme into an assay which can detect the presence of specific nucleic acid sequences. By splitting the catalytic core of a deoxyribozyme, the DNAzyme loses its catalytic activity. However reformation of the core

between the two strands can restore the molecules enzymatic abilities (see Figure 1b) [112, 113]. The DNAzyme can also be altered so that the transition from inactive to active state only occurs in the presence of a specific analyte sequence. The DNAzyme can then cleave a substrate molecule, which may vary for fluorescent (see figure 1c) or visual output. This reaction can be measured by a plate reader or spectrophotometer [114, 115]. Astonishingly, the binary DNAzyme assay has been shown to be specific enough to differentiate between two targets that differ by only a single SNP [116, 117].

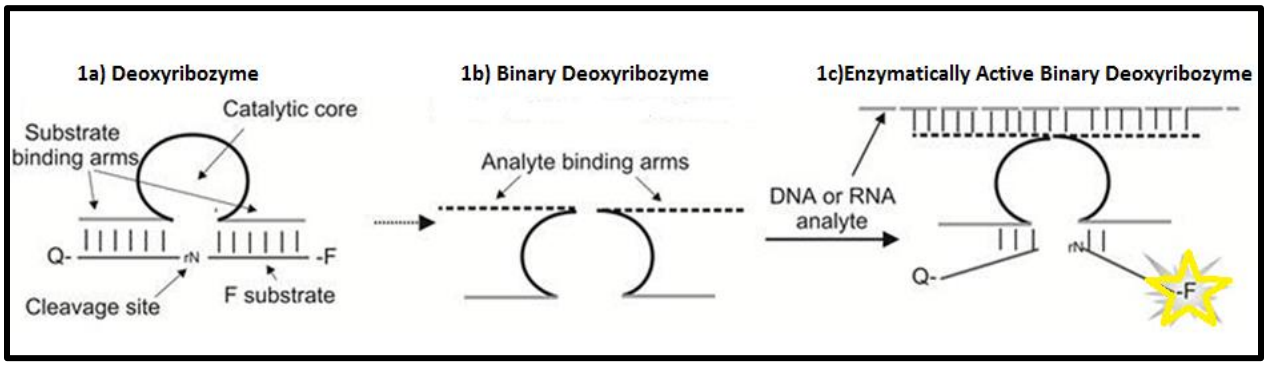

Figure 1: Development of the Binary Deoxyribozyme. 1a) A standard Deoxyribozyme consists of a catalytic core and substrate binding arms. 1b) a Binary Deoxyribozyme consists of 2 separate strands each with analyte binding arms complimentary to a target sequence. 1c) When bound to a target analyte, the catalytic core is able to be reformed, allowing binding of the substrate binding arms and cleavage of a substrate molecule to release a florescent output.

Designing a binary DNAzyme sensor assay first requires basic knowledge of the binary DNAzyme structure. A fully assembled binary DNAzyme consists of three primary structures, the analyte binding arms, the catalytic core, and the substrate binding arms [116]. Each sensor strand therefore contains one analyte binding arm, one substrate binding arm and one half of the catalytic core (see figure 2).

The substrate binding arm sequence is designed to be complimentary to the sequence of the substrate molecule. The substrate molecule for the fluorescent assay is a synthetic DNA-

RNA hybrid oligonucleotide analyte, containing a ribonucleotide directly in the center of the molecule allowing for cleavage [114]. Other modifications include a fluorophore and a quencher at the 5' and 3' of the molecule respectively [114]. The specific substrate utilized is known as Mzf Substrate and consists of a FAM fluorophore and BHQ quencher.

The catalytic core of the DNAzyme is responsible for facilitating the actual enzymatic reaction the molecule catalyzes. Variants of the sequence of the catalytic core exist, resulting in a higher or lower catalytic rate [105]. The use of different catalytic cores also alters the chemical reactions that the DNAzyme can perform such as bond cleavage or peroxidase-like activity [118].

Similar to the design of the substrate binding arm, the analyte binding arm is designed to be complimentary to the target sequence being analyzed, such as sequences found in the rRNA of *Mtb* [119]. Designing these arms requires previous knowledge of the sequence being analyzed, and it is these arms which are redesigned to create a new binary DNAzyme assay.

In a fluorescent binary DNAzyme assay, the complementarity of the analyte and the analyte binding arms results in both strands hybridizing to the analyte. This consequently leads to the reformation of the catalytic core and allows the substrate binding arms to bind to the substrate molecule. The newly functional binary DNAzyme then catalyzes the cleavage of the substrate molecule which separates the fluorophore and the quencher. This liberated fluorophore can be quantified as fluorescent units using a plate reader or spectrophotometer. The full structure of a binary DNAzyme complexed with its matched analyte is shown in below.

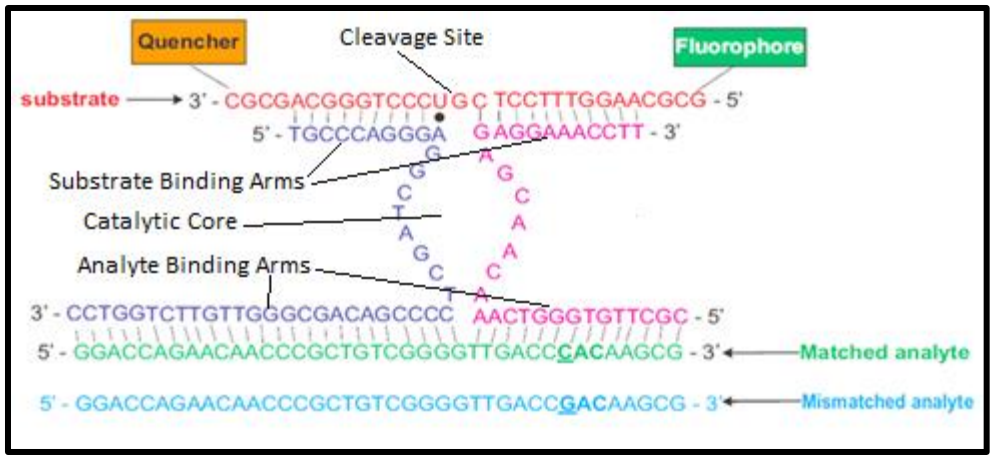

Figure 2: Structure of a Binary DNAzyme. All binary DNAzymes consist of substrate binding arms which have a complimentary sequence to the substrate molecule, a catalytic core which catalyzes cleavage of the substrate molecule, and analyte binding arms which are complimentary in sequence to a target analyte. Designing a new sensor pair only requires redesigning the analyte binding arms.

The binary DNAzyme has many advantages over conventional molecular diagnostic techniques. Because substrate molecule is universal between all fluorescent DNAzyme assays, designing a new assay does not require redesigning or reordering the substrate molecule. Altering the probes for a totally new assay simply requires redesigning the two strands which bind to the new target analyte [120]. These sensors have been implemented in the detection of bacterial nucleic acids, however they can be redesigned to detect other nucleic acid based infectious agents such as viruses [120]. Furthermore, detection of human SNPs with binary DNAzymes is also possible [121]. Some other advantages of the DNAzyme sensor assay are that they do not require trained personnel, and lack the requirement of an amplification step when targeting genes present in high copy number.

Methods and Materials

Binary DNAzyme Assay

General Binary DNAzyme Assay

After designing a binary DNAzyme, the sensors are tested with a matched and mismatched analyte to assure their correct design. Each fluorescent assay also has two intrinsic controls built in to analyze both stability of the substrate molecule and the background signal produced from unbound sensor strands. The first control consists of 200nM of the Mzf substrate molecule in Col buffer. Col buffer consists of 50Mm HEPES, 50mM MgCl2, 20mM KCl, 120mM NaCl, 0.03% Triton X-100 and 1% DMSO. This buffer supplies the DNAzyme with all the essential cations required for catalytic function. Correct concentration of both the substrate and ions within the Col buffer is insured by utilizing 30µl of a 2X Col buffer, adding the substrate molecule, and diluting the sample with dH_2O up to the standard total reaction volume of 60µl.

Next, a master mix is made consisting of the two sensor strands, substrate, Col 2X buffer, and dH_2O. Total volume necessary for each test sample along with a background sample and one additional sample is first determined ($ReactionVol_{Total}$). The master mix consists of Col 2X buffer at a volume of one half this total reaction volume ($ReactionVol_{Total}/2$). Each sensor strand is then added to the master mix at a concentration of 15nM along with the substrate molecule at a concentration of 200nM. To determine dH_2O added to the master mix, the total volume is multiplied by .75 to give the total master mix volume ($MasterMixVol = .75 \times ReactionVol_{Total}$). dH_2O is then added to the master mix to bring to total volume up to this calculated value.

Each sample tube then receives 45µl of the Master mix. 15µl of dH_2O is added to the background tube, forming the second internal standard for comparison. The sample tubes receive the matched or mismatched analyte at a concentration of 1nM and dH_2O is then added to bring the total volume of the tubes up to 60µl.

The samples along with both the substrate and background control are placed in a 384 well black bottom plate and incubated at 54°C for 1 hour in a hybridization oven. The fluorescent intensity is then measured in a Synergy 4 Biotek Plate Reader. Detection using the plate reader has been optimized, with best results using Filter Based detection with excitation and emission wavelengths of 485nm and 520nm respectively and a read height of 1mm. Each fluorescent signal is then divided by the signal of the background control, allowing the measurements to be interpreted as a signal to background ratio (S/B).

Specificity and Limit of Detection

Assays were designed for *Mabs* with comparable TB sensors being tested by other members of the lab. Two varieties of the standard binary DNAzyme assay include a specificity study along with a limit of detection study. The goal of the specificity study goal is to compare the signal to background ratio obtained by the matched analyte with that of the mismatched analyte in the form of a *specificity value*. Three different types of analyte were routinely tested: synthetic oligonucleotides, total RNA, and PCR amplicons. 1nM of synthetic analyte is utilized while 200ng of total RNA is used. PCR amplicon concentration is difficult to determine without purifying out any residual dNTPs, so 15µl of the PCR product is routinely utilized. When testing total RNA as analyte, an additional binary DNAzyme assay is commonly performed using sensors complimentary for the sequence of the mismatched analyte to assure integrity of the

RNA. The specificity study is performed in triplicate in order to generate averages along with standard deviation.

The limit of detection (LOD) study utilizes many samples of serial dilution of matched analyte in order to form a linear graph which correlates S/B ratios with target analyte concentration. The target analyte is serial diluted in regular intervals to produce many samples of decreasing concentration. A DNAzyme assay is then performed in triplicate on the various analyte samples, and a scatterplot is generated comparing S/B ratio (Y) with analyte concentration (X). The LOD value is the concentration of analyte which produces a signal to background ratio equal to that of the background sample (no analyte) plus three times the background samples standard deviation [122]. The LOD test can be performed on both synthetic analytes as well as RNA.

Visual Identification

Visual detection is accomplished using a second DNAzyme, known as IPDZ, as the substrate molecule in what is known as a DNAzyme cascade [118]. The binary DNAzyme facilitates cleavage of a ribonucleotide separating an inhibitory component of the second DNAzyme, liberating the catalytic PDz fragment [123]. This DNAzyme folds into a G-Quadruplex formation and, in the presence of a hemin cofactor, has peroxidase activity which can be detected via the oxidation of the colorimetric substrate DAB in the presence of H_2O_2 (see figure 4). The resulting molecule changes the solution from clear to brown, allowing visual detection or absorbance measurement [117, 123].

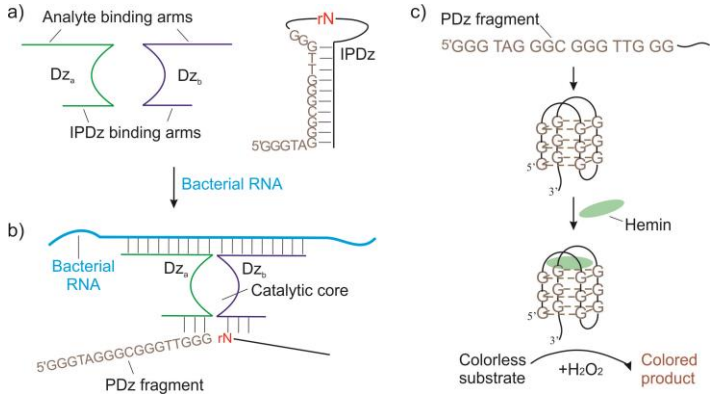

Figure 3: Visual Reaction Cascade. A) A visual binary DNAzyme assay requires the two binary DNAzyme strands, each consisting of an analyte binding arm, substrate binding arm, and half of a catalytic core. The assay also requires an inhibited peroxidase like DNAzyme (IPDz). B) In the presence of a specific analyte, the binary DNAzyme is able to bind and reform the catalytic core, allowing cleavage of the inhibitory fragment from the IPDz to form PDz. C) Liberated PDz adopts a G-quadruplex like structure and in the presence of hemin can catalyze an oxidoreductive reaction to produce a colored product.

A visual binary DNAzyme assay consists of the same setup and controls as a fluorescent binary DNAzyme assay with a few minor alterations. The initial substrate molecule, IPDZ, is used at a 1µM concentration while each sensor strand is used at 100nM. The assay is then incubated at 50°C for 1 hour. Afterwards, 2µl of a 1:1 mixture of DAB and hemin at 25µM is added to each tube followed by 1µl of H_2O_2. After 30 minutes of incubation at room temperature the absorbance is measured at 500nm in a spectrophotometer. Specificity tests were accomplished utilizing 100nM of analyte.

SNP Analysis

Specific design parameters must be applied to binary DNAzymes intended for the detection of single SNPS. These sensors are constructed with a long arm and a short arm, with the short arm containing the discriminatory SNP. This methodology allows the strand containing the long arm to be utilized for both detection of both wild type and mutant variations. Because the SNP-containing arm is shorter than the nondiscriminatory arm its lower annealing temperature allows for more sensitive mismatch discrimination [117].

Our approach to DNAzyme-based SNP detection uses an algorithm for strain typing developed by Homolka *et al.* that defined the minimal number of SNPs required for high resolution lineage assignment identification (see figure 3) [104]. As proof-of-principle for the utility of our technology for strain typing we designed sensors to analyze the G→C mutation of the 455th residue of the *Rv0557* gene which allows the determination of Haarlem vs. NonHaarlem lineage (See appendix).

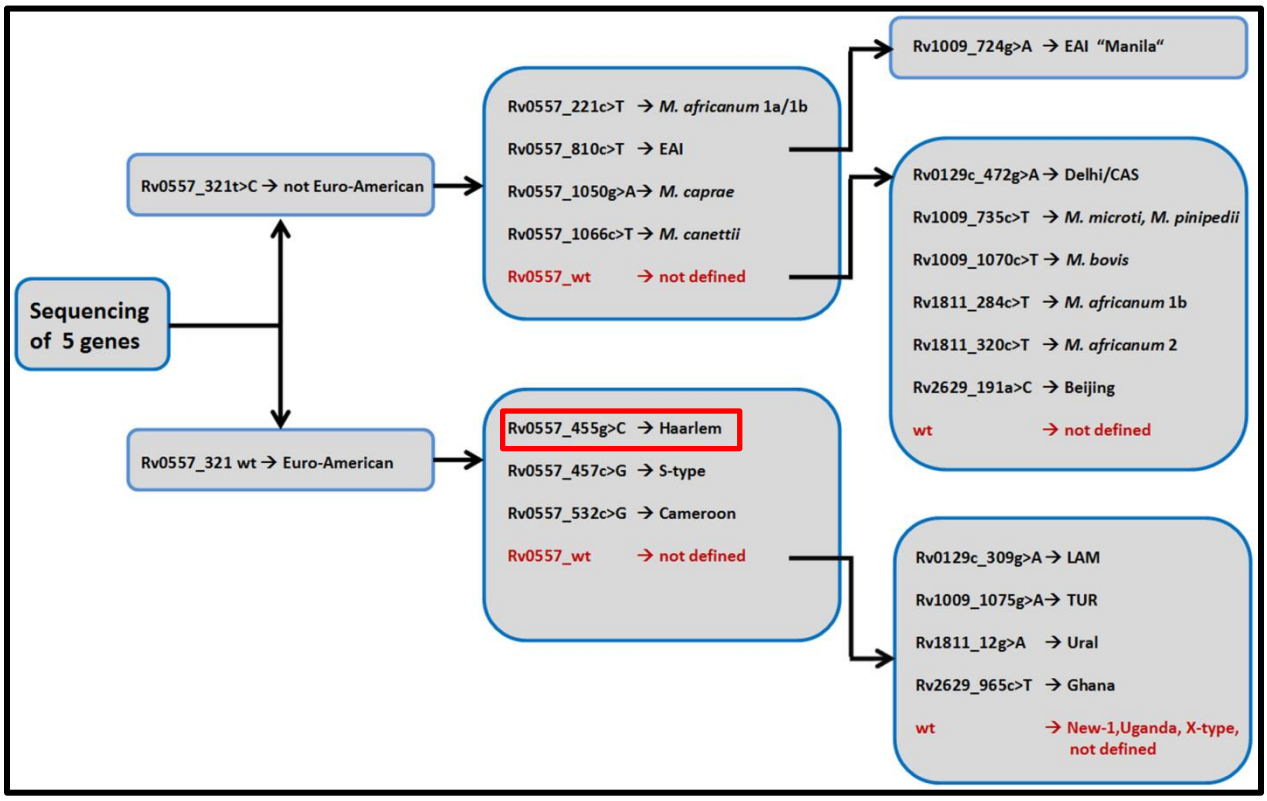

Figure 4: Strain Typing SNPs. Homolka *et al.* were able to determine the minimum number of SNPs needed to be sequenced in order to effectively characterize strain lineage. This was accomplished computationally after multiple rounds of whole genome sequencing on different strains of *Mtb*.

As a proof of principle experiment to characterize drug resistance, we also designed sensors to detect the most common drug resistance mutations within the *rpoB* gene, a C526G mutation which confers Rif resistance (See table).

Target Amplification

Analysis of SNPs in genes which are present in low copy number requires PCR amplification in order for efficient detection. PCR amplification of a target sequence requires two primers which bind both upstream and downstream to the target sequence respectively. The sequence of the forward primer is identical the sense strand while the reverse primer is the reverse complement. A standard PCR reaction includes 1µl DNA at approximately 20ng/µl concentration, 1µl each forward and reverse primers at 100ng/µl concentration, .25µl of Phusion DNA polymerase at 2000U/ml concentration, 5µl of HF Buffer, 16.25µl of dH_2O, and .5µl of dNTPs at 10mM concentration. The PCR reaction sample is then placed into a Philisa thermal cycler which can quickly cycle through various temperatures, allowing the DNA to denature, the primers to anneal to the separated DNA, and the polymerase to extend the primers and form a new strand of DNA. Initially the sample is heated to 98°C for 3 minutes, followed by 50 cycles of 98°C for 10 seconds, 67°C for 15 seconds, and 72°C for 10 seconds. The temperature is then held at 72°C for 30 seconds and cooled to room temperature. In order to determine drug resistance, PCR primers were designed and utilized for the *rpoB* gene in an effort to characterize Rif resistance (see appendix).

It has been hypothesized that a binary DNAzyme is more sensitive utilizing single stranded DNA (ssDNA) and an analyte compared to double stranded DNA due to needing to compete with the antisense strand for hybridization. One method of generating ssDNA is asymmetric PCR. In asymmetric PCR, one primer is decreased in concentration significantly (1:40). This results in exponential increase in DNA during the initial cycles until the limiting primer runs out. After depleting the limiting primer, extension of only the excess primer causes

the linear production of ssDNA. LATE-PCR, OR Linear After The Exponential PCR, accounts for the fact that decreasing the concentration of one of the primers decreases the effective annealing temperature, leading to ineffective amplification [124]. LATE-PCR addresses this problem via altered design of the limiting primer, where it is extended by 2 to 3 nucleotides. This raises the effective annealing temperature of the limiting primer, allowing for an increase in output [125]. Both LATE-PCR and symmetric PCR primers were designed for the amplification of a small fragment of the 23S rRNA gene in *Mtb* (See Appendix).

Results and Discussion

Mabs Detection

Fluorescent Detection

Fluorescent sensors to detect *Mabs* were developed which targeted regions of the *23S rRNA*, *16S rRNA* and *tmRNA* genes [126]. The sequences selected based upon multiple sequence alignments were homologous among different strains of *Mabs*, allowing detection of all strains. The regions selected however were variable between different species of Mycobacterium, allowing for efficient species differentiation from Mycobacterium tuberculosis complex. Because the genes chosen are present in high copy number, PCR amplification was not needed.

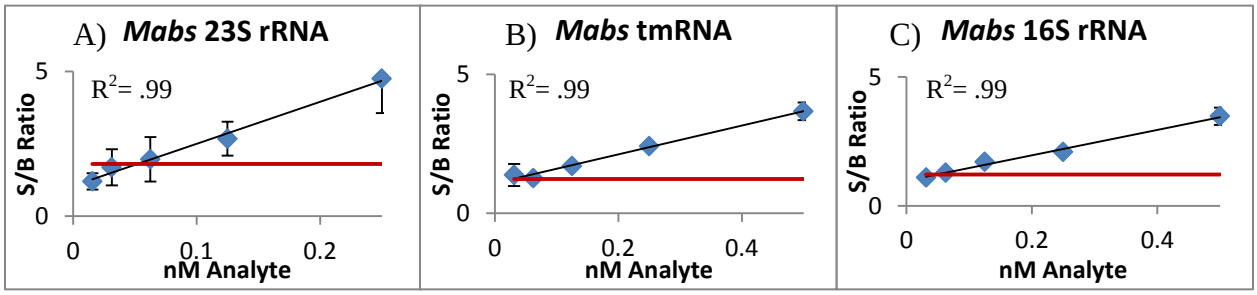

Figure 5: *Mabs* 23S rRNA, tmRNA and 16S rRNA Limit of Detection using Synthetic Analyte. A) B) and C) represent signal to background ratio of fluorescence at various concentrations of synthetic analyte for *Mabs* 23S rRNA, tmRNA and 16S rRNA respectively. The red line represents limit of detection. Each graph represents the averages of 3 separate experiments.

Limit of detection studies were completed in triplicate using both synthetic analyte and total RNA using a range of 0.5nM (0.25nM for 23S rRNA) to 0.015625nM for synthetic analyte and 100ng to 6.25ng for total RNA. The red line indicates the S/B Ratio for the limit of detection. The limit of detection using synthetic analyte was determined to be 48.18pM, 51.76pM and 28.54pM for 16S rRNA, 23S rRNA and tmRNA sensors respectively. These numbers are similar to that of sensors previously designed to detect 16S rRNA and 23S rRNA in *Mtb*, indicating effective detection. Limit of detection also indicates quality of the sensors, as the

analyte concentration used to determine the value is the same, allowing comparison between the sensors.

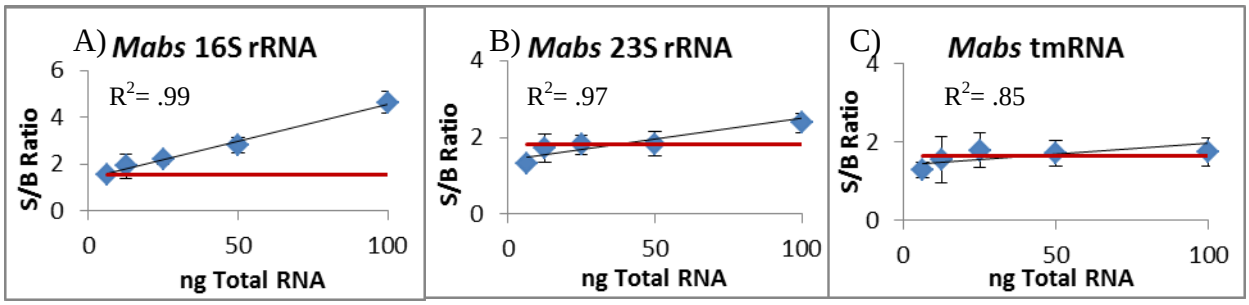

Figure 6: *Mabs* 23S rRNA, tmRNA and 16S rRNA Limit of Detection using Total RNA. A) B) and C) represent limit of detection studies utilizing total RNA at various concentrations for *Mabs* 16S rRNA, 23S rRNA and tmRNA respectively. The red line represents the limit of detection. Each graph represents the average of 3 independent experiments.

The limit of detection using total RNA was determined to be 4.34ng, 37.42ng and 41.07ng for 16S rRNA, 23S rRNA and tmRNA sensors respectively. The differences in these values are not able to indicate differences in quality of the sensors, as the limit of detection using RNA is also influenced by the total copy number in the total RNA. The results of this study demonstrate the potential to detect mycobacterial RNA directly without the need for amplification.

In addition to limit of detection, the specificity ratio of the 23S rRNA and tmRNA *Mabs* sensors were determined by comparing the S/B of the perfect match analyte (*Mabs*) with the mismatch analyte (*Mtb*). A high specificity ratio assures that the sensors will be able to accurately differentiate between different species of Mycobacterium, preventing the possibility of a false positive.

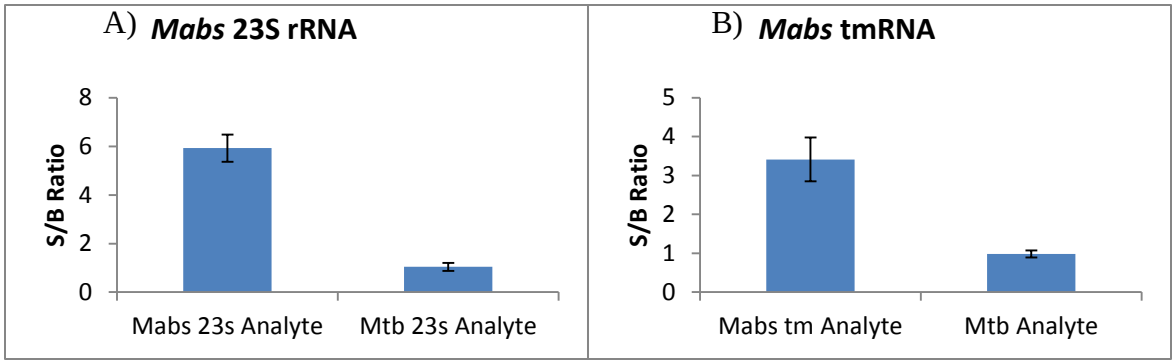

Figure 7: *Mabs* 23S rRNA and tmRNA Specificity using Synthetic Analyte. A) and B) represent specificity of Mabs 23S rRNA and tmRNA sensors respectively. Sensors were tested in triplicate with synthetic matched (*Mabs*) and mismatched (*Mtb*) analyte.

The *specificity ratios* of the 23S rRNA and tmRNA sensors using synthetic analyte were determined to be 5.69 and 3.46 respectively. The matched analyte for the 23S rRNA *Mabs* sensors gave a S/B Ratio of 5.69 while the mismatched analyte only produced a signal of 1.04. The difference in the specificity ratios between the 23S rRNA and the tmRNA sensors is explained by the lower overall signal the matched analyte produced, with a S/B Ratio of 3.46 compared to 1.25 produced by the mismatched analyte. These results show that our sensors were accurately able to discriminate against different genetic regions between *Mabs* and *Mtb*, providing the ability to differentiate between the two mycobacteria.

Visual Detection

Visual sensors to detect *Mabs* which target the *23S rRNA*, *16S rRNA* and *tmRNA* genes were also developed and tested for their limit of detection and specificity ratio. These sensors were easily designed, utilizing the same analyte binding arms and catalytic core portion as the fluorescent sensors, and only required altering the substrate binding arms. In order to quantify the color change due to the colorimetric reaction, absorbance was measured at 500nm. A positive readout can also be obtained by simply observing a change in color, indicating the

presence of a matched analyte. The ability to detect this color change by the naked eye is necessary for POC diagnostics as it eliminates the need for expensive equipment to analyze the results.

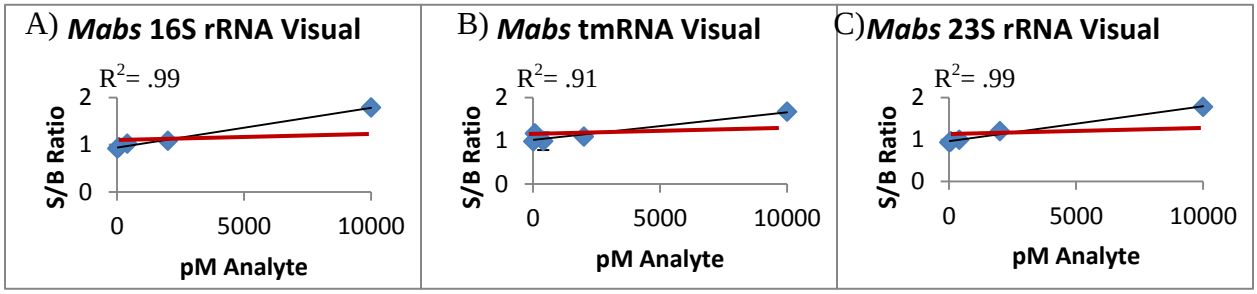

Figure 8: *Mabs* 23S rRNA, tmRNA and 16S rRNA Visual Limit of Detection using Synthetic Analyte. A) B) and C) represent signal to background of absorbance (500nm) using synthetic analyte at various concentrations for *Mabs* 16S rRNA, tmRNA, and 23S rRNA respectively. Each graph represents the average of 3 separate experiments. Visual detection was also confirmed via a color change. The red line represents limit of detection.

Limit of detection studies were completed in triplicate using synthetic analyte at a concentration range of 10nM to 16pM. The limit of detection was determined to be 946pM, 948pM and 2624pM for 16S rRNA, tmRNA and 23S rRNA visual sensors respectively. This is not surprising, as the limit of detection of visual sensors are commonly seem to be at least ten times higher than that of their fluorescent counterpart [123].

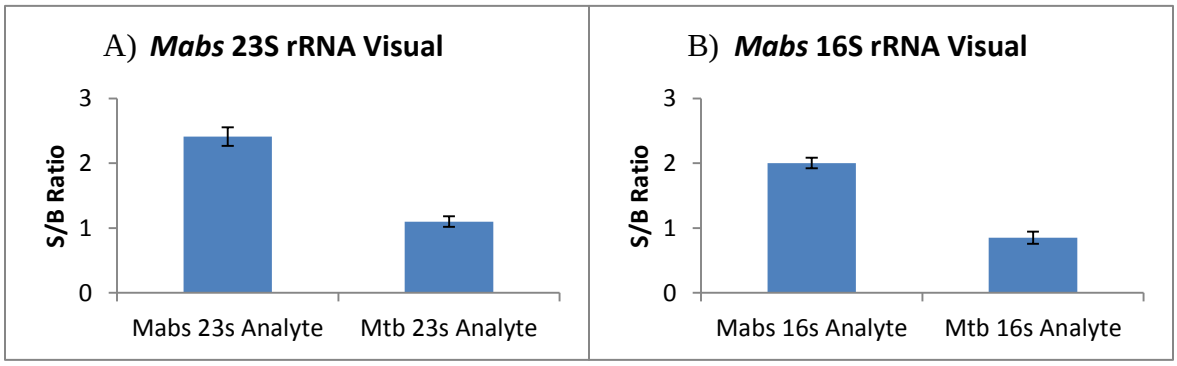

Figure 9: *Mabs* 23S rRNA and tmRNA Visual Sensors Specificity. A) and B) represent specificity of *Mabs* visual 23S rRNA and 16S rRNA sensors respectively. Absorbance (500nm) of each sensor was tested 3 independent times with matched (*Mabs*) and mismatched (*Mtb*) synthetic analyte. The difference in specificity was also observed visually.

The specificity ratios were experimentally determined in triplicate using synthetic analyte, with the 23S rRNA and 16S rRNA visual sensors scoring a 2.19 and 2.35 respectively. The S/B Ratio for the matched analyte of the *Mabs* 23S rRNA sensors was 2.41 while the S/B ratio for the mismatched analyte was 1.10. The S/B Ratio for the matched analyte of the *Mabs* 16S rRNA sensors was 2.00 while the S/B ratio for the mismatched analyte was 0.85. The S/B Ratios of visual sensors are also commonly seen to be lower than that of fluorescent sensors. This is due to the lower signal range of visual assay versus fluorescence. S/B Ratios using a matched analyte on a visual sensor's typically reaches a maximum of about 2.5.

Amplification

Symmetric vs. LATE-PCR

Analyzing genes present in low copy number requires the use of an application technique. We compared traditional (symmetric) PCR to LATE-PCR (asymmetric) in order to determine the most efficient amplification technique for using a binary DNAzyme. LATE-PCR generates ssDNA due to decreasing the concentration of an extended limiting primer. Primers were designed to amplify a small fragment of the *23S rRNA* gene in *Mtb*. The PCR products were then run on an agarose gel and stained with Gel Red to image. The appearance of a second band in the LATE-PCR sample indicates the production of ssDNA. The PCR product was then subjected to a binary DNAzyme assay using *Mtb* 23S rRNA sensors.

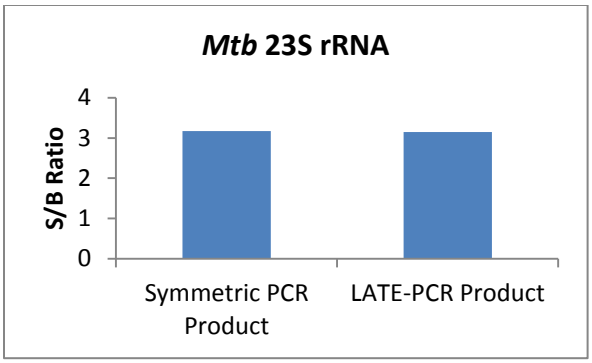

Figure 10: Symmetric vs. LATE-PCR using 23S *Mtb* rRNA Sensors. 15µl of symmetric PCR and LATE-PCR product amplifying a fragment of *Mtb* 23S rRNA was tested using a *Mtb* 23S rRNA fluorescent sensor.

Surprisingly, the S/B ratio using symmetric PCR compared to LATE-PCR was nearly identical, which was contrary to our working hypothesis. Figure 11 represents the data from a single repetition; however the results were replicated with similar results in multiple independent experiments. The reason the generation of ssDNA does not affect the S/B ratio is still unknown, although one possible explanation is because the generation of ssDNA is linear, the total concentration of amplicon, single and double stranded, is significantly less than that of symmetric PCR. Another explanation could be that due the small size of the amplicon, the binary DNAzyme sensors are able to efficiently outcompete the antisense strand from binding to the target. The excess concentration of the sensors may also increase the signal when using a PCR amplicon as an analyte.

Cycle Optimization

In order to determine optimal cycle number, primers were designed and used to amplify a 250 base pair fragment of the *rpoB* gene of wild type *Mtb* using 30x, 40x, and 50x cycles. This was confirmed using agarose gel electrophoresis. A binary DNAzyme assay using *rpoB* wild type sensors was then completed to determine how cycle number affects S/B ratio.

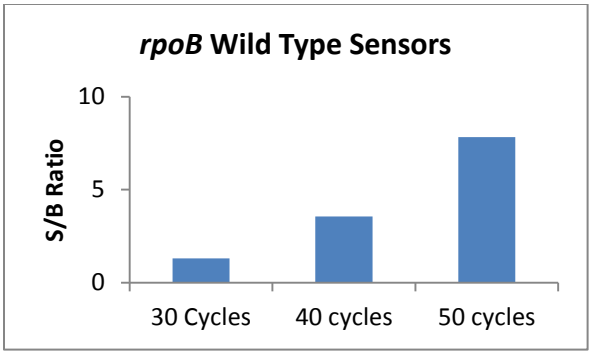

Figure 11: *rpoB* Wild Type Sensors on Determining Optimum PCR Cycle Number. 15µl of standard PCR product amplifying a fragment of the *Mtb rpoB* gene at various cycle numbers was tested using wild type *rpoB* sensors.

The maximum S/B of 7.82 achieved with 50 cycles led us to select this as the optimum cycle number. Figure 10 indicates a single rep, however further experiments have confirmed our findings.

SNP Analysis

Strain Typing

We next attempted to utilize the ability of binary DNAzyme sensors to detect SNPs for molecular strain typing of *Mtb*. Based on a SNP-based typing algorithm proposed by Homolka et al (2012), we designed sensors to detect the WT and mutant variants of the G455C SNP in the gene r*v0557* which distinguishes the Haarlem from non-Haarlem lineage. In this initial proof of principle experiment, the two sensor pairs were tested against synthetic analytes representing the Haarlem and non-Haarlem sequence at this site.

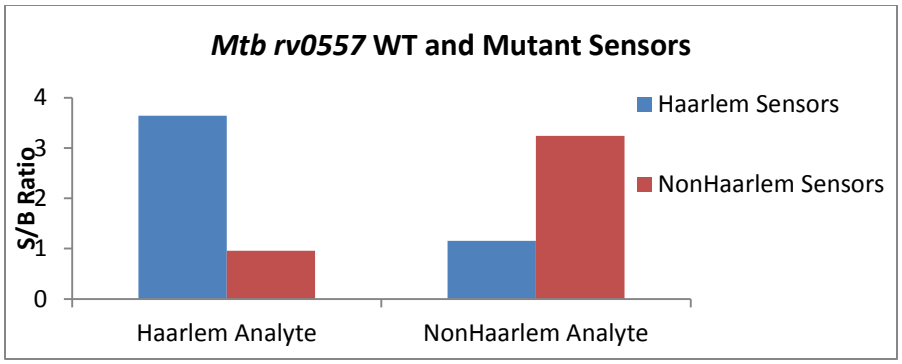

Figure 12: Haarlem Vs. NonHaarlem Sensor Specificity. Both Haarlem and NonHaarlem fluorescent sensors were tested using their matched and mismatched analyte respectively. These sensors targeted a G→C mutation of the 455th nucleotide of the *rv0557* gene.

Both Haarlem and NonHaarlem sensors showed excellent specificity against synthetic analytes, with a *specificity ratio* of 3.14 and 3.39 in one rep. These results combined with our data showing binary DNAzyme detection of PCR amplicons suggest a PCR-based binary DNAzyme assay would be effective for strain typing *Mtb* from these two distinct lineages.

Drug Resistance Analysis

Sensors to detect wild type *rpoB* gene were developed and tested against wild type and mutant analyte. The sensors designed detect wild type sequence which would indicate susceptibility to Rif. This means that a loss of signal would indicate the presence of a drug resistant mutant. A wild type sensor was chosen because the loss of signal would occur for any mutation to the section of the *rpoB* gene being amplified, as opposed to detecting just one of three possible nucleotides which would confer drug resistance.

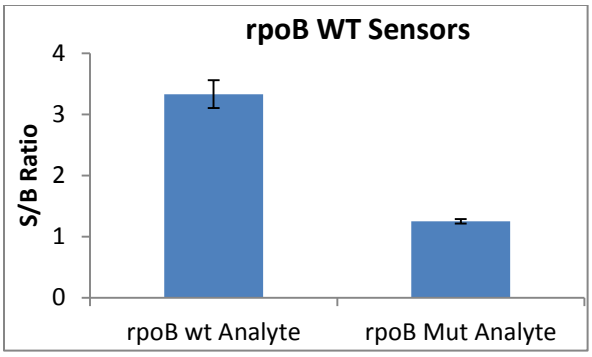

Figure 13: Specificity of *rpoB* Sensors using Wild-Type and Mutant Synthetic Analyte. Specificity was determined using both matched (wild type) and mismatched (drug resistant) synthetic analyte. This was performed in triplicate.

Using synthetic analyte the wild type sensors showed an acceptable level of specificity, with a specificity ratio of 2.67. The matched analyte resulted in a S/B ratio of 3.33 while a mismatched analyte containing the C526G mutation resulted in a signal of 1.25. Because the *rpoB* gene is present in low copy number, primers were designed to amplify a 250 base pair fragment of the gene containing the 81 base pair hot spot region. The specificity of the wild type binary DNAzyme sensors were then determined using the PCR sample produced from both wild type and Rif resistant BCG DNA. This DNA was generated from artificially selecting an rpoB mutation in a strand of BCG within the lab, while the sequence was confirmed by independent sequencing.

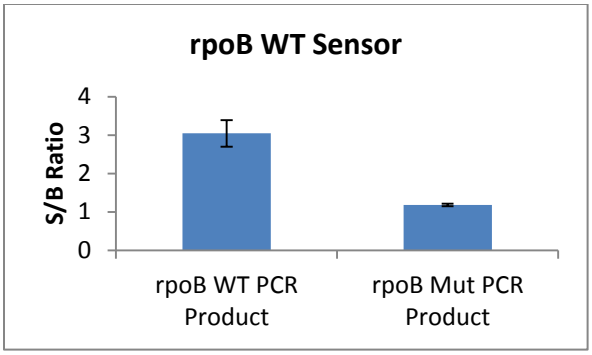

Figure 14: *rpoB* Wild-Type Sensor Specificity using PCR Product. Specificity was determined using *rpoB* wild type sensors after amplification using CDC1551 WT DNA and artificially selected Rif resistant BCG DNA. This was performed in triplicate.

The specificity value of the wild type rpoB sensors using PCR product decreased to a 2.58, still indicating efficient discrimination between drug susceptible and drug resistant variation of *Mtb*.

Future Works

With these experiments, we have shown that binary DNAzyme sensors can be designed to efficiently detect and characterize different species of pathogenic mycobacterium. The development of visual sensors also results in the generation of an assay which can translate into POC diagnostics due to the ease of reading the output. We are currently testing out other peroxidase substrate molecules which can potentially be used instead of DAB and produce an even greater absorbance output. We also plan on testing other methods of detection such as a 384 well filter bottom plate, allowing the colored substrate molecule to be condensed into the well and result in a stronger visual output. This would also be an alternative to using a plate reader, eliminating the need for expensive equipment.

The characterization of the *rpoB* gene is vital in determining drug resistance or susceptibility towards Rif, a first line drug used to treat *Mtb*. Because of the development of MDR-Tb and XDR-Tb, it is necessary to design and test sensors which can characterize resistance to additional antibiotics including Inh and FQs. We have currently developed primers to amplify regions of the *inhA, katG* and *gyrA* genes, which can be combined with the *rpoB* primers in one PCR reaction. This should allow the generation of a multiplex PCR product which can be subjected to multiple binary DNAzyme assays to elucidate drug resistance characterization to many different drugs.

In an effort to eliminate the need for a thermal cycler, we are currently testing isothermal amplification techniques such as Loop Mediated Isothermal Amplification (LAMP). LAMP utilizes 6 primers that recognize 6 distinct regions of DNA in a target sequence and results in the generation of concatemeric stem loop DNA structures, all linked together with the same target

sequence [127]. The elimination of a thermal cycler greatly increases the possibility that an assay combining LAMP and the binary DNAzyme can directly translate into an efficient POC diagnostics.

Summary

The detection and differentiation of pathogenic mycobacterium is essential in diagnosis and treatment of the disease. Early diagnosis results in higher survival rates among patients, however current methods of diagnostics can be unreliable and expensive. We have shown that sensors developed to detect *Mabs* are able to detect low quantities of total RNA and synthetic analyte and are efficiently able to differentiate between *Mabs* and *Mtb*. We also designed sensors to detect *Mabs* with a visual output, eliminating the need for a plate reader and successfully advancing the development of a POC diagnostic test for the detection of pathogenic mycobacteria.

SNP characterization in *Mtb* allows for the elucidation of a variety of molecular characteristics of the individual strain, such as lineage and drug resistance profile. We were able to successfully design sensors which were sensitive enough to differentiate between two strains with a single residue difference. These sensors efficiently discriminate between two different lineages of *Mtb*, allowing for effective strain typing without the need for whole genome sequencing.

Characterization of drug resistance is of vital importance when treating an *Mtb* infection, and determining drug susceptibility is vital for determining the course of treatment. We were able to show that binary DNAzyme sensors were able to efficiently differentiate between Rif susceptible and Rif resistant *Mtb*. We also showed that PCR amplification may be necessary in characterizing SNPs found in genes which are present in low copy number, and the symmetric PCR is effectively able to amplify a large enough quantity of amplicons to reliably detect with DNAzyme sensor. We also showed that a binary DNAzyme sensor is still efficiently able to

differentiate between the different SNPs when using PCR amplicons as the analyte. Overall, these experiments make ways to facilitate better treatment of TB by providing a novel platform for POC detection. We have shown that the binary DNAzyme platform is a promising technology to fill this urgent need.

Appendix

Appendix

Fluorescent Sensors

Substrate: MzF	5’- AAGGTT-FAM-TCCTCguCCCTGGGCA-BHQ	
Sensor	**Adaptor Strand A**	**Adaptor Strand B (WT specific)**
rpoB Sensor	k2 (A1_rpoB_526_WT)	k1 (B1_rpoB_526_WT)
	5’-TGCCCAGGGAGGCTAGCTCCGACAGCGGGTTGTTCTGGTCC	5’-CGCTTGTGGGTCAAACAACGAGAGGAAACCTT
Analyte	**Sequence**	
k3 (rpoB_526_WT_CAC)	5’- GGACCAGAACAACCCGCTGTCGGGGTTGACCCACAAGCG	
MT K3 -DN (rpoB_526_MT_TAC)	5’- GGACCAGAACAACCCGCTGTCGGGGTTGACCTACAAGCG	
Sensor	**Adaptor Strand A**	**Adaptor Strand B**
Mabs 23S Sensor	23S_M_abs_1a	23S_M_abs_1b
	5’-TGCCCAGGGAGGCTAGCTCAAGTCCGGTCACCAAGCGAGGT	5’-GTGCGACCCACTCACGCTTGGGCCGACAACGAGAGGAAACCTT
Analyte	**Sequence**	
23S_M_abs-1	5’- ACCTCGCTTGGTGACCGGACTTGCTCCGTGAGCTGAACGAGGTCGTCGCAG	

Sensor	**Adaptor Strand A**	**Adaptor Strand B**
Mabs 16S Sensor	16S_M_abs_1a	16S_M_abs_1b
	5'- TGCCCAGGGAGGCTAGCT TGAAGTGTGTGGTCCTATCCGG	5'- CAAAGCTTTGCACCACTCACCA ACAACGAGAGGAAACCTT
Analyte	**Sequence**	
16S_M_abs-1	5'-CCG GAT AGG ACC ACA CAC TTC ATG GTG AGT GGT GCA AAG CTT TTG	
Sensor	**Adaptor Strand A**	**Adaptor Strand B**
Mabs TM Sensor	tm_Mabs_1a	tm_Mabs_1b
	5'- TGCCCAGGGAGGCTAGCT GGCCGAAGCCACCGGTTGATCCCG	5'- GTTTGATGTCCCGGTGGCCTCCGC ACAACGAGAGGAAACCTT
Analyte	**Sequence**	
tm_M_ab	5'-GCGGGATCAACCGGTGGCTTCGGCCGCGGAGGCCACCGGGACATCAAC	
Sensor	**Adaptor Strand A**	**Adaptor Strand B**
Mtb 23s Sensor	23S_Mtb_1a	23S_Mtb_1b
	5'- TCGCCAGGGAGGCTAGCTCAAGGGTGTTCACGAT ATTCCGGT	5'- CTGCGACCGGATCCCGCTCCCACCGACAACGAGAG GAAACCTT
Analyte	**Sequence**	
23S_Mtb-1	5'-ACCGGAATATCGTGAACACCCTTGCGGTGGGAGCGGGATCCGGTCGCAG	

Sensor	**Adaptor Strand A**	**Adaptor Strand B**
Rv0557 Haarlem Sensor	nonHaarlem_455_G>T A1_fl	Haarlem_455_T B1 fl
	5'-CTGTGGACGGCGCCAGAGTGCGGTCACAACGAGAGGAAACCTT	5'-TGCCCAGGGAGGCTAGCTAGCCAGGGGATGCAA
Analyte	**Sequence**	
Haarlem_455_C	5'- CCACTTGCATCCCCTGGCTGACCGCACTCTGGCGCCGTCCACAGCGAC	
Sensor	**Adaptor Strand A**	**Adaptor Strand B**
Rv0557 Non-Haarlem Sensor	nonHaarlem_455_G>T A1_fl	nonHaarlem_455_G_WT_B1_fl
	5'-CTGTGGACGGCGCCAGAGTGCGGTCACAACGAGAGGAAACCTT	5'-TGCCCAGGGAGGCTAGCTAGCCAGGCGATGCCA
Analyte	**Sequence**	
nonHaarlem_455_G-WT	5'- CCACTTGCATCGCCTGGCTGACCGCACTCTGGCGCCGTCCACAGCGAC	

Visual Sensors

Substrate: IPDZ	5‘- GGG TAG GGC GGG TTG GGT TC rG rU CC ATG AGC AACTCG CCC	
Sensor	**Adaptor Strand A**	**Adaptor Strand B**
Mabs 23S Visual Sensor	23S_M_abs_1a_vis	23S_M_abs_1b_vis
	5’-GTTGCTCATGGAGGCTAGCTCAAGTCCGGTCACCAAGCGA	5’-GTTTGATGTCCCGGTGGCCTCCGCACAACGAGAACCCAACC
Sensor	**Adaptor Strand A**	**Adaptor Strand B**
Mabs 16S Visual Sensor	16S_M_abs_1a_vis	16S_M_abs_1b_vis
	5’-GTTGCTCATGGAGGCTAGCTTGAAGTGTGTGGTCCTATCCGG	5’-CAAAGCTTTGCACCACTCACCAACAACGAGAACCCAACC
Sensor	**Adaptor Strand A**	**Adaptor Strand B**
Mabs TM Visual Sensor	tm_Mabs_1a_vis	tm_Mabs_1b_vis
	5’-GTTGCTCATGGAGGCTAGCTGGCCGAAGCCACCGGTTGATCCCG	5’-GTTTGATGTCCCGGTGGCCTCCGCACAACGAGAACCCAACC

Note: Same Analytes are used as Fluorescent Assay

Primers

Primers		
Locus	**Primer**	**Sequence**
rpoB	rpoB_sym200_F	5’- GTCGCCGCGATCAAGGAGTT
	rpoB_sym200_R	5’- CCCTCAGGGGTTTCGATCGGG

Locus	**Primer**	**Sequence**
23S rRNA Mtb	TB23S_aPCR_F	5’- GAACTCGGCAAAATGCCCCCGTAA
	Tb23s_aPCR_R (For LATE-PCR)	5’- TTCGTGCAGGTCGGAACTTACCCCGACAAGG
	23SR_sym	5’- TTCGTGCAGGTCGGAACTTACCCG

Reference List

1. WHO, *Global Tuberculosis Report 2013. Geneva, Switzerland.* 2013.
2. Cambau, E. and M. Drancourt, *Steps towards the discovery of Mycobacterium tuberculosis by Robert Koch, 1882.* Clin Microbiol Infect, 2014. **20**(3): p. 196-201.
3. Donoghue, H.D., et al., *Tuberculosis in Dr Granville's mummy: a molecular re-examination of the earliest known Egyptian mummy to be scientifically examined and given a medical diagnosis.* Proc Biol Sci, 2010. **277**(1678): p. 51-6.
4. Barksdale, L. and K.S. Kim, *Mycobacterium.* Bacteriol Rev, 1977. **41**(1): p. 217-372.
5. Barry, C.E., 3rd, et al., *Mycolic acids: structure, biosynthesis and physiological functions.* Prog Lipid Res, 1998. **37**(2-3): p. 143-79.
6. Korf, J., et al., *The Mycobacterium tuberculosis cell wall component mycolic acid elicits pathogen-associated host innate immune responses.* Eur J Immunol, 2005. **35**(3): p. 890-900.
7. Daffe, M. and P. Draper, *The envelope layers of mycobacteria with reference to their pathogenicity.* Adv Microb Physiol, 1998. **39**: p. 131-203.
8. Orme, I.M., *A new unifying theory of the pathogenesis of tuberculosis.* Tuberculosis (Edinb), 2014. **94**(1): p. 8-14.
9. Ferrari, G., et al., *A coat protein on phagosomes involved in the intracellular survival of mycobacteria.* Cell, 1999. **97**(4): p. 435-47.
10. Rhoades, E.R., A.A. Frank, and I.M. Orme, *Progression of chronic pulmonary tuberculosis in mice aerogenically infected with virulent Mycobacterium tuberculosis.* Tuber Lung Dis, 1997. **78**(1): p. 57-66.

11. McElvania Tekippe, E., et al., *Granuloma formation and host defense in chronic Mycobacterium tuberculosis infection requires PYCARD/ASC but not NLRP3 or caspase-1.* PLoS One, 2010. **5**(8): p. e12320.

12. Cardona, P.J., et al., *Evolution of granulomas in lungs of mice infected aerogenically with Mycobacterium tuberculosis.* Scand J Immunol, 2000. **52**(2): p. 156-63.

13. Miner, M.D., et al., *Role of cholesterol in Mycobacterium tuberculosis infection.* Indian J Exp Biol, 2009. **47**(6): p. 407-11.

14. Basaraba, R.J., *Experimental tuberculosis: the role of comparative pathology in the discovery of improved tuberculosis treatment strategies.* Tuberculosis (Edinb), 2008. **88 Suppl 1**: p. S35-47.

15. Reiley, W.W., et al., *Maintenance of peripheral T cell responses during Mycobacterium tuberculosis infection.* J Immunol, 2012. **189**(9): p. 4451-8.

16. Salgame, P., *Host innate and Th1 responses and the bacterial factors that control Mycobacterium tuberculosis infection.* Curr Opin Immunol, 2005. **17**(4): p. 374-80.

17. Moore, M. and J.B. Frerichs, *An unusual acid-fast infection of the knee with subcutaneous, abscess-like lesions of the gluteal region; report of a case with a study of the organism, Mycobacterium abscessus, n. sp.* J Invest Dermatol, 1953. **20**(2): p. 133-69.

18. Kusunoki, S. and T. Ezaki, *Proposal of Mycobacterium peregrinum sp. nov., nom. rev., and elevation of Mycobacterium chelonae subsp. abscessus (Kubica et al.) to species status: Mycobacterium abscessus comb. nov.* Int J Syst Bacteriol, 1992. **42**(2): p. 240-5.

19. Horsburgh, C.R., Jr., *Epidemiology of disease caused by nontuberculous mycobacteria.* Semin Respir Infect, 1996. **11**(4): p. 244-51.

20. Chadha, R., et al., *An outbreak of post-surgical wound infections due to Mycobacterium abscessus.* Pediatr Surg Int, 1998. **13**(5-6): p. 406-10.

21. Zhibang, Y., et al., *Large-scale outbreak of infection with Mycobacterium chelonae subsp. abscessus after penicillin injection.* J Clin Microbiol, 2002. **40**(7): p. 2626-8.

22. Griffith, D.E., W.M. Girard, and R.J. Wallace, Jr., *Clinical features of pulmonary disease caused by rapidly growing mycobacteria. An analysis of 154 patients.* Am Rev Respir Dis, 1993. **147**(5): p. 1271-8.

23. Fauroux, B., et al., *Mycobacterial lung disease in cystic fibrosis: a prospective study.* Pediatr Infect Dis J, 1997. **16**(4): p. 354-8.

24. Griffith, D.E., et al., *An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases.* Am J Respir Crit Care Med, 2007. **175**(4): p. 367-416.

25. Bryant, J.M., et al., *Whole-genome sequencing to identify transmission of Mycobacterium abscessus between patients with cystic fibrosis: a retrospective cohort study.* Lancet, 2013. **381**(9877): p. 1551-60.

26. Wallace, R.J., Jr., et al., *Spectrum of disease due to rapidly growing mycobacteria.* Rev Infect Dis, 1983. **5**(4): p. 657-79.

27. Vogels, M.T., et al., *Cutaneous infection due to Mycobacterium abscessus. A case report.* Acta Derm Venereol, 1997. **77**(3): p. 222-4.

28. Winthrop, K.L., et al., *Pulmonary nontuberculous mycobacterial disease prevalence and clinical features: an emerging public health disease.* Am J Respir Crit Care Med, 2010. **182**(7): p. 977-82.

29. Jeon, K., et al., *Antibiotic treatment of Mycobacterium abscessus lung disease: a retrospective analysis of 65 patients.* Am J Respir Crit Care Med, 2009. **180**(9): p. 896-902.

30. Griffith, D.E. and R.J. Wallace, Jr., *New developments in the treatment of nontuberculous mycobacterial (NTM) disease.* Semin Respir Infect, 1996. **11**(4): p. 301-10.

31. Esteban, J., et al., *Detection of lfrA and tap efflux pump genes among clinical isolates of non-pigmented rapidly growing mycobacteria.* Int J Antimicrob Agents, 2009. **34**(5): p. 454-6.

32. Nessar, R., et al., *Mycobacterium abscessus: a new antibiotic nightmare.* J Antimicrob Chemother, 2012. **67**(4): p. 810-8.

33. Zhang, Y., et al., *Role of acid pH and deficient efflux of pyrazinoic acid in unique susceptibility of Mycobacterium tuberculosis to pyrazinamide.* J Bacteriol, 1999. **181**(7): p. 2044-9.

34. Ainsa, J.A., et al., *Molecular cloning and characterization of Tap, a putative multidrug efflux pump present in Mycobacterium fortuitum and Mycobacterium tuberculosis.* J Bacteriol, 1998. **180**(22): p. 5836-43.

35. Choudhuri, B.S., et al., *Overexpression and functional characterization of an ABC (ATP-binding cassette) transporter encoded by the genes drrA and drrB of Mycobacterium tuberculosis.* Biochem J, 2002. **367**(Pt 1): p. 279-85.

36. Siddiqi, N., et al., *Mycobacterium tuberculosis isolate with a distinct genomic identity overexpresses a tap-like efflux pump.* Infection, 2004. **32**(2): p. 109-11.

37. Silva, P.E., et al., *Characterization of P55, a multidrug efflux pump in Mycobacterium bovis and Mycobacterium tuberculosis.* Antimicrob Agents Chemother, 2001. **45**(3): p. 800-4.

38. Reisner, B.S., A.M. Gatson, and G.L. Woods, *Evaluation of mycobacteria growth indicator tubes for susceptibility testing of Mycobacterium tuberculosis to isoniazid and rifampin.* Diagn Microbiol Infect Dis, 1995. **22**(4): p. 325-9.

39. Miller, L.P., J.T. Crawford, and T.M. Shinnick, *The rpoB gene of Mycobacterium tuberculosis.* Antimicrob Agents Chemother, 1994. **38**(4): p. 805-11.

40. Anthony, R.M., et al., *Acquisition of rifabutin resistance by a rifampicin resistant mutant of Mycobacterium tuberculosis involves an unusual spectrum of mutations and elevated frequency.* Ann Clin Microbiol Antimicrob, 2005. **4**: p. 9.

41. Karahan, Z.C., et al., *Sequence analysis of rpoB mutations in rifampin-resistant clinical Mycobacterium tuberculosis isolates from Turkey.* Microb Drug Resist, 2004. **10**(4): p. 325-33.

42. Quemard, A., et al., *Enzymatic characterization of the target for isoniazid in Mycobacterium tuberculosis.* Biochemistry, 1995. **34**(26): p. 8235-41.

43. Banerjee, A., et al., *inhA, a gene encoding a target for isoniazid and ethionamide in Mycobacterium tuberculosis.* Science, 1994. **263**(5144): p. 227-30.

44. Rozwarski, D.A., et al., *Modification of the NADH of the isoniazid target (InhA) from Mycobacterium tuberculosis.* Science, 1998. **279**(5347): p. 98-102.

45. Basso, L.A., et al., *Mechanisms of isoniazid resistance in Mycobacterium tuberculosis: enzymatic characterization of enoyl reductase mutants identified in isoniazid-resistant clinical isolates.* J Infect Dis, 1998. **178**(3): p. 769-75.

46. Mdluli, K., et al., *Biochemical and genetic data suggest that InhA is not the primary target for activated isoniazid in Mycobacterium tuberculosis.* J Infect Dis, 1996. **174**(5): p. 1085-90.

47. Fenner, L., et al., *Effect of mutation and genetic background on drug resistance in Mycobacterium tuberculosis.* Antimicrob Agents Chemother, 2012. **56**(6): p. 3047-53.

48. Cockerill, F.R., 3rd, et al., *Rapid identification of a point mutation of the Mycobacterium tuberculosis catalase-peroxidase (katG) gene associated with isoniazid resistance.* J Infect Dis, 1995. **171**(1): p. 240-5.

49. Heym, B., et al., *Missense mutations in the catalase-peroxidase gene, katG, are associated with isoniazid resistance in Mycobacterium tuberculosis.* Mol Microbiol, 1995. **15**(2): p. 235-45.

50. Zhang, Y., et al., *The catalase-peroxidase gene and isoniazid resistance of Mycobacterium tuberculosis.* Nature, 1992. **358**(6387): p. 591-3.

51. Rouse, D.A., et al., *Characterization of the katG and inhA genes of isoniazid-resistant clinical isolates of Mycobacterium tuberculosis.* Antimicrob Agents Chemother, 1995. **39**(11): p. 2472-7.

52. Drlica, K., et al., *Fluoroquinolone action in mycobacteria: similarity with effects in Escherichia coli and detection by cell lysate viscosity.* Antimicrob Agents Chemother, 1996. **40**(7): p. 1594-9.

53. Bryskier, A., *Fluoroquinolones: mechanisms of action and resistance.* Int J Antimicrob Agents, 1993. **2**(3): p. 151-83.

54. Alangaden, G.J., et al., *Characterization of fluoroquinolone-resistant mutant strains of Mycobacterium tuberculosis selected in the laboratory and isolated from patients.* Antimicrob Agents Chemother, 1995. **39**(8): p. 1700-3.

55. Piddock, L.J., *Mechanisms of resistance to fluoroquinolones: state-of-the-art 1992-1994.* Drugs, 1995. **49 Suppl 2**: p. 29-35.

56. Palomino, J.C., *Nonconventional and new methods in the diagnosis of tuberculosis: feasibility and applicability in the field.* Eur Respir J, 2005. **26**(2): p. 339-50.

57. Greenwood, N. and H. Fox, *A comparison of methods for staining tubercle bacilli in histological sections.* J Clin Pathol, 1973. **26**(4): p. 253-7.

58. WHO, *Laboratory services in tuberculosis control.Part II. Microscopy. Geneva, Switzerland.* 1998.

59. Chen, P., et al., *A highly efficient Ziehl-Neelsen stain: identifying de novo intracellular Mycobacterium tuberculosis and improving detection of extracellular M. tuberculosis in cerebrospinal fluid.* J Clin Microbiol, 2012. **50**(4): p. 1166-70.

60. Rigouts, L., *Clinical practice: diagnosis of childhood tuberculosis.* Eur J Pediatr, 2009. **168**(11): p. 1285-90.

61. Dinic, L., et al., *Sputum smear concentration may misidentify acid-fast bacilli as Mycobacterium tuberculosis in HIV-infected patients.* J Acquir Immune Defic Syndr, 2013. **63**(2): p. 168-77.

62. Lagier, J., et al., *Current and Past Strategies for Bacterial Culture in Clinical Microbiology.* Clin Microbiol Rev, 2015. **28**(1): p. 208-236.

63. Palaci, M., et al., *Evaluation of mycobacteria growth indicator tube for recovery and drug susceptibility testing of Mycobacterium tuberculosis isolates from respiratory specimens.* J Clin Microbiol, 1996. **34**(3): p. 762-4.

64. Margileth, A.M., *The use of purified protein derivative mycobacterial skin test antigens in children and adolescents: purified protein derivative skin test results correlated with mycobacterial isolates.* Pediatr Infect Dis, 1983. **2**(3): p. 225-31.

65. Rosenthal, S.R., *Standardization and efficacy of BCG vaccination against tuberculosis; twenty year study: a critical evaluation.* J Am Med Assoc, 1955. **157**(10): p. 801-7.

66. Friedland, I.R., *The booster effect with repeat tuberculin testing in children and its relationship to BCG vaccination.* S Afr Med J, 1990. **77**(8): p. 387-9.

67. Singla, M., et al., *BCG skin reaction in Mantoux-negative healthy children.* BMC Infect Dis, 2005. **5**: p. 19.

68. Shankar, P., et al., *Identification of Mycobacterium tuberculosis by polymerase chain reaction.* Lancet, 1990. **335**(8686): p. 423.

69. Abed, Y., C. Bollet, and P. de Micco, *Identification and strain differentiation of Mycobacterium species on the basis of DNA 16S-23S spacer region polymorphism.* Res Microbiol, 1995. **146**(5): p. 405-13.

70. Cho, S.N. and P.J. Brennan, *Tuberculosis: diagnostics.* Tuberculosis (Edinb), 2007. **87 Suppl 1**: p. S14-7.

71. Yue, J., et al., *Detection of rifampin-resistant Mycobacterium tuberculosis strains by using a specialized oligonucleotide microarray.* Diagn Microbiol Infect Dis, 2004. **48**(1): p. 47-54.

72. Williams, D.L., T.P. Gillis, and W.G. Dupree, *Ethanol fixation of sputum sediments for DNA-based detection of Mycobacterium tuberculosis.* J Clin Microbiol, 1995. **33**(6): p. 1558-61.

73. Helb, D., et al., *Rapid detection of Mycobacterium tuberculosis and rifampin resistance by use of on-demand, near-patient technology.* J Clin Microbiol, 2010. **48**(1): p. 229-37.

74. Malbruny, B., et al., *Rapid and efficient detection of Mycobacterium tuberculosis in respiratory and non-respiratory samples.* Int J Tuberc Lung Dis, 2011. **15**(4): p. 553-5.

75. Bodmer, T. and A. Strohle, *Diagnosing pulmonary tuberculosis with the Xpert MTB/RIF test.* J Vis Exp, 2012(62): p. e3547.

76. Morris, K., *Xpert TB diagnostic highlights gap in point-of-care pipeline.* Lancet Infect Dis, 2010. **10**(11): p. 742-3.

77. Coscolla, M. and S. Gagneux, *Consequences of genomic diversity in Mycobacterium tuberculosis.* Semin Immunol, 2014. **26**(6): p. 431-444.

78. Zink, A.R. and A.G. Nerlich, *Molecular strain identification of the Mycobacterium tuberculosis complex in archival tissue samples.* J Clin Pathol, 2004. **57**(11): p. 1185-92.

79. Kato-Maeda, M., et al., *The nature and consequence of genetic variability within Mycobacterium tuberculosis.* J Clin Invest, 2001. **107**(5): p. 533-7.

80. Rohde, K.H., et al., *Linking the transcriptional profiles and the physiological states of Mycobacterium tuberculosis during an extended intracellular infection.* PLoS Pathog, 2012. **8**(6): p. e1002769.

81. Ford, C., et al., *Mycobacterium tuberculosis--heterogeneity revealed through whole genome sequencing.* Tuberculosis (Edinb), 2012. **92**(3): p. 194-201.

82. Ford, C.B., et al., *Mycobacterium tuberculosis mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis.* Nat Genet, 2013. **45**(7): p. 784-90.

83. Provvedi, R., G. Palu, and R. Manganelli, *Use of DNA microarrays to study global patterns of gene expression.* Methods Mol Biol, 2009. **465**: p. 95-110.

84. Lopes, J.S., et al., *SNP typing reveals similarity in Mycobacterium tuberculosis genetic diversity between Portugal and Northeast Brazil.* Infect Genet Evol, 2013. **18**: p. 238-46.

85. Ahmed, N., et al., *Genome sequence based, comparative analysis of the fluorescent amplified fragment length polymorphisms (FAFLP) of tubercle bacilli from seals provides molecular evidence for a new species within the Mycobacterium tuberculosis complex.* Infect Genet Evol, 2003. **2**(3): p. 193-9.

86. Cohn, D.L. and R.J. O'Brien, *The use of restriction fragment length polymorphism (RFLP) analysis for epidemiological studies of tuberculosis in developing countries.* Int J Tuberc Lung Dis, 1998. **2**(1): p. 16-26.

87. van Embden, J.D., et al., *Strain identification of Mycobacterium tuberculosis by DNA fingerprinting: recommendations for a standardized methodology.* J Clin Microbiol, 1993. **31**(2): p. 406-9.

88. Kamerbeek, J., et al., *Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology.* J Clin Microbiol, 1997. **35**(4): p. 907-14.

89. Kanduma, E., T.D. McHugh, and S.H. Gillespie, *Molecular methods for Mycobacterium tuberculosis strain typing: a users guide.* J Appl Microbiol, 2003. **94**(5): p. 781-91.

90. Botelho, A., et al., *Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs) Analysis of Members of the Mycobacterium tuberculosis Complex.* Methods Mol Biol, 2015. **1247**: p. 373-89.

91. Goyal, M., et al., *Differentiation of Mycobacterium tuberculosis isolates by spoligotyping and IS6110 restriction fragment length polymorphism.* J Clin Microbiol, 1997. **35**(3): p. 647-51.

92. Magdalena, J., et al., *Identification of a new DNA region specific for members of Mycobacterium tuberculosis complex.* J Clin Microbiol, 1998. **36**(4): p. 937-43.

93. Gibson, A., et al., *Can 15-locus mycobacterial interspersed repetitive unit-variable-number tandem repeat analysis provide insight into the evolution of Mycobacterium tuberculosis?* Appl Environ Microbiol, 2005. **71**(12): p. 8207-13.

94. Allix-Beguec, C., et al., *Evaluation and strategy for use of MIRU-VNTRplus, a multifunctional database for online analysis of genotyping data and phylogenetic identification of Mycobacterium tuberculosis complex isolates.* J Clin Microbiol, 2008. **46**(8): p. 2692-9.

95. Majeed, A.A., et al., *AmpliBASE MT: a Mycobacterium tuberculosis diversity knowledgebase.* Bioinformatics, 2004. **20**(6): p. 989-92.

96. Comas, I., et al., *Genotyping of genetically monomorphic bacteria: DNA sequencing in Mycobacterium tuberculosis highlights the limitations of current methodologies.* PLoS One, 2009. **4**(11): p. e7815.

97. Tsolaki, A.G., et al., *Genomic deletions classify the Beijing/W strains as a distinct genetic lineage of Mycobacterium tuberculosis.* J Clin Microbiol, 2005. **43**(7): p. 3185-91.

98. Alland, D., et al., *Role of large sequence polymorphisms (LSPs) in generating genomic diversity among clinical isolates of Mycobacterium tuberculosis and the utility of LSPs in phylogenetic analysis.* J Clin Microbiol, 2007. **45**(1): p. 39-46.

99. Roetzer, A., et al., *Whole genome sequencing versus traditional genotyping for investigation of a Mycobacterium tuberculosis outbreak: a longitudinal molecular epidemiological study.* PLoS Med, 2013. **10**(2): p. e1001387.

100. Fleischmann, R.D., et al., *Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains.* J Bacteriol, 2002. **184**(19): p. 5479-90.

101. Fleischmann, R.D., et al., *Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.* Science, 1995. **269**(5223): p. 496-512.

102. Filliol, I., et al., *Global phylogeny of Mycobacterium tuberculosis based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set.* J Bacteriol, 2006. **188**(2): p. 759-72.

103. Schurch, A.C., et al., *SNP/RD typing of Mycobacterium tuberculosis Beijing strains reveals local and worldwide disseminated clonal complexes.* PLoS One, 2011. **6**(12): p. e28365.

104. Homolka, S., et al., *High resolution discrimination of clinical Mycobacterium tuberculosis complex strains based on single nucleotide polymorphisms.* PLoS One, 2012. **7**(7): p. e39855.

105. Santoro, S.W. and G.F. Joyce, *A general purpose RNA-cleaving DNA enzyme.* Proc Natl Acad Sci U S A, 1997. **94**(9): p. 4262-6.

106. Appaiahgari, M.B. and S. Vrati, *DNAzyme-mediated inhibition of Japanese encephalitis virus replication in mouse brain.* Mol Ther, 2007. **15**(9): p. 1593-9.

107. Stojanovic, M.N., et al., *Deoxyribozyme-based ligase logic gates and their initial circuits.* J Am Chem Soc, 2005. **127**(19): p. 6914-5.

108. Ruble, B.K., et al., *Mismatch Discrimination and Efficient Photomodulation with Split 10-23 DNAzymes.* Inorganica Chim Acta, 2012. **380**: p. 386-391.

109. Breaker, R.R. and G.F. Joyce, *A DNA enzyme that cleaves RNA.* Chem Biol, 1994. **1**(4): p. 223-9.

110. Joyce, G.F., *In vitro evolution of nucleic acids.* Curr Opin Struct Biol, 1994. **4**: p. 331-6.

111. Hansen, M.H., et al., *A yoctoliter-scale DNA reactor for small-molecule evolution.* J Am Chem Soc, 2009. **131**(3): p. 1322-7.

112. Kolpashchikov, D.M., *Binary malachite green aptamer for fluorescent detection of nucleic acids.* J Am Chem Soc, 2005. **127**(36): p. 12442-3.

113. Bichenkova, E.V., et al., *Target-assembled tandem oligonucleotide systems based on exciplexes for detecting DNA mismatches and single nucleotide polymorphisms.* Biochem Biophys Res Commun, 2005. **332**(4): p. 956-64.

114. Kolpashchikov, D.M., *A binary deoxyribozyme for nucleic acid analysis.* Chembiochem, 2007. **8**(17): p. 2039-42.

115. Gerasimova, Y.V. and D.M. Kolpashchikov, *Nucleic acid detection using MNAzymes.* Chem Biol, 2010. **17**(2): p. 104-6.

116. Kolpashchikov, D.M., *A binary DNA probe for highly specific nucleic Acid recognition.* J Am Chem Soc, 2006. **128**(32): p. 10625-8.

117. Kolpashchikov, D.M., *Split DNA enzyme for visual single nucleotide polymorphism typing.* J Am Chem Soc, 2008. **130**(10): p. 2934-5.

118. Li, T., et al., *Multifunctional G-quadruplex aptamers and their application to protein detection.* Chemistry, 2009. **15**(4): p. 1036-42.

119. Gerasimova, Y.V. and D.M. Kolpashchikov, *Detection of bacterial 16S rRNA using a molecular beacon-based X sensor.* Biosens Bioelectron, 2013. **41**: p. 386-90.

120. Kolpashchikov, D.M., *Binary probes for nucleic acid analysis.* Chem Rev, 2010. **110**(8): p. 4709-23.

121. Gerasimova, Y.V., J. Ballantyne, and D.M. Kolpashchikov, *Detection of SNP-containing human DNA sequences using a split sensor with a universal molecular beacon reporter.* Methods Mol Biol, 2013. **1039**: p. 69-80.

122. Gerasimova, Y.V., E. Cornett, and D.M. Kolpashchikov, *RNA-cleaving deoxyribozyme sensor for nucleic acid analysis: the limit of detection.* Chembiochem, 2010. **11**(6): p. 811-7, 729.

123. Gerasimova, Y.V., et al., *Deoxyribozyme cascade for visual detection of bacterial RNA.* Chembiochem, 2013. **14**(16): p. 2087-90.

124. Pierce, K.E., et al., *Linear-After-The-Exponential (LATE)-PCR: primer design criteria for high yields of specific single-stranded DNA and improved real-time detection.* Proc Natl Acad Sci U S A, 2005. **102**(24): p. 8609-14.

125. Sanchez, J.A., et al., *Linear-after-the-exponential (LATE)-PCR: an advanced method of asymmetric PCR and its uses in quantitative real-time analysis.* Proc Natl Acad Sci U S A, 2004. **101**(7): p. 1933-8.

126. Mignard, S. and J.P. Flandrois, *Identification of Mycobacterium using the EF-Tu encoding (tuf) gene and the tmRNA encoding (ssrA) gene.* J Med Microbiol, 2007. **56**(Pt 8): p. 1033-41.

127. Notomi, T., et al., *Loop-mediated isothermal amplification of DNA.* Nucleic Acids Res, 2000. **28**(12): p. E63.